

OIE

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/897,844A

DATE: 11/23/2001
 TIME: 10:57:17

Input Set : N:\Crf3\RULE60\09897844A.RAW
 Output Set: N:\CRF3\11232001\I897844A.raw

ENTERED

1 <110> APPLICANT: Cox III, George Norbert
 2 Case, Casey Christopher
 3 Eisenberg, Stephen P.
 4 Jarvis, Eric Edward
 5 Spratt, Sharon Kaye
 6 Sangamo Biosciences, Inc.
 7 <120> TITLE OF INVENTION: Regulation of Endogenous Gene Expression in Cells
 8 Using Zinc Finger Proteins
 9 <130> FILE REFERENCE: 019496-002200US
 10 <140> CURRENT APPLICATION NUMBER: 09/897,844A
 11 <141> CURRENT FILING DATE: 2001-07-02
 12 <150> PRIOR APPLICATION NUMBER: US/09/229,037
 13 <151> PRIOR FILING DATE: 1999-01-12
 14 <160> NUMBER OF SEQ ID NOS: 40
 15 <170> SOFTWARE: PatentIn Ver. 2.0
 17 <210> SEQ ID NO: 1
 18 <211> LENGTH: 25
 19 <212> TYPE: PRT
 20 <213> ORGANISM: Artificial Sequence
 21 <220> FEATURE:
 22 <223> OTHER INFORMATION: Description of Artificial Sequence:exemplary motif
 23 of C2H2 class of zinc finger proteins (ZFP)
 24 <220> FEATURE:
 25 <221> NAME/KEY: MOD_RES
 26 <222> LOCATION: (2)..(3)
 27 <223> OTHER INFORMATION: Xaa = any amino acid
 28 <220> FEATURE:
 29 <221> NAME/KEY: MOD_RES
 30 <222> LOCATION: (4)..(5)
 31 <223> OTHER INFORMATION: Xaa = any amino acid, may be present or absent
 32 <220> FEATURE:
 33 <221> NAME/KEY: MOD_RES
 34 <222> LOCATION: (7)..(18)
 35 <223> OTHER INFORMATION: Xaa = any amino acid
 36 <220> FEATURE:
 37 <221> NAME/KEY: MOD_RES
 38 <222> LOCATION: (20)..(22)
 39 <223> OTHER INFORMATION: Xaa = any amino acid
 40 <220> FEATURE:
 41 <221> NAME/KEY: MOD_RES
 42 <222> LOCATION: (23)..(24)
 43 <223> OTHER INFORMATION: Xaa = any amino acid, may be present or absent
 44 <400> SEQUENCE: 1
 W--> 45 Cys Xaa Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 46 1 5 10 15
 W--> 47 Xaa Xaa His Xaa Xaa Xaa Xaa His
 48 20 25

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/897,844A

DATE: 11/23/2001
 TIME: 10:57:17

Input Set : N:\Crf3\RULE60\09897844A.RAW
 Output Set: N:\CRF3\11232001\I897844A.raw

50 <210> SEQ ID NO: 2
 51 <211> LENGTH: 10
 52 <212> TYPE: DNA
 53 <213> ORGANISM: Artificial Sequence
 54 <220> FEATURE:
 55 <223> OTHER INFORMATION: Description of Artificial Sequence:ZFP target site
 56 with two overlapping D-able subsites
 57 <220> FEATURE:
 58 <221> NAME/KEY: modified_base
 59 <222> LOCATION: (1)..(2)
 60 <223> OTHER INFORMATION: n = g,a,c or t
 61 <220> FEATURE:
 62 <221> NAME/KEY: modified_base
 63 <222> LOCATION: (5)
 64 <223> OTHER INFORMATION: n = g,a,c or t
 65 <220> FEATURE:
 66 <221> NAME/KEY: modified_base
 67 <222> LOCATION: (8)
 68 <223> OTHER INFORMATION: n = g,a,c or t
 69 <220> FEATURE:
 70 <221> NAME/KEY: modified_base
 71 <222> LOCATION: (9)
 72 <223> OTHER INFORMATION: n = a,c or t; if g, then position 10 cannot be g
 73 or t
 74 <220> FEATURE:
 75 <221> NAME/KEY: modified_base
 76 <222> LOCATION: (10)
 77 <223> OTHER INFORMATION: n = a or c; if g or t, then position 9 cannot be g
 78 <400> SEQUENCE: 2 10

W--> 79 nngkngknnn

81 <210> SEQ ID NO: 3
 82 <211> LENGTH: 10
 83 <212> TYPE: DNA
 84 <213> ORGANISM: Artificial Sequence
 85 <220> FEATURE:
 86 <223> OTHER INFORMATION: Description of Artificial Sequence:ZFP target site
 87 with three overlapping D-able subsites
 88 <220> FEATURE:
 89 <221> NAME/KEY: modified_base
 90 <222> LOCATION: (1)..(2)
 91 <223> OTHER INFORMATION: n = g,a,c or t
 92 <220> FEATURE:
 93 <221> NAME/KEY: modified_base
 94 <222> LOCATION: (5)
 95 <223> OTHER INFORMATION: n = g,a,c or t
 96 <220> FEATURE:
 97 <221> NAME/KEY: modified_base
 98 <222> LOCATION: (8)
 99 <223> OTHER INFORMATION: n = g,a,c or t

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/09/897,844A

DATE: 11/23/2001
 TIME: 10:57:17

Input Set : N:\Crf3\RULE60\09897844A.RAW
 Output Set: N:\CRF3\11232001\I897844A.raw

100 <400> SEQUENCE: 3
 W--> 101 nngkngkngk 10
 103 <210> SEQ ID NO: 4
 104 <211> LENGTH: 5
 105 <212> TYPE: PRT
 106 <213> ORGANISM: Artificial Sequence
 107 <220> FEATURE:
 108 <223> OTHER INFORMATION: Description of Artificial Sequence:linker
 109 <400> SEQUENCE: 4
 110 Asp Gly Gly Gly Ser
 111 1 5
 113 <210> SEQ ID NO: 5
 114 <211> LENGTH: 5
 115 <212> TYPE: PRT
 116 <213> ORGANISM: Artificial Sequence
 117 <220> FEATURE:
 118 <223> OTHER INFORMATION: Description of Artificial Sequence:linker
 119 <400> SEQUENCE: 5
 120 Thr Gly Glu Lys Pro
 121 1 5
 123 <210> SEQ ID NO: 6
 124 <211> LENGTH: 9
 125 <212> TYPE: PRT
 126 <213> ORGANISM: Artificial Sequence
 127 <220> FEATURE:
 128 <223> OTHER INFORMATION: Description of Artificial Sequence:linker
 129 <400> SEQUENCE: 6
 130 Leu Arg Gln Lys Asp Gly Glu Arg Pro
 131 1 5
 133 <210> SEQ ID NO: 7
 134 <211> LENGTH: 4
 135 <212> TYPE: PRT
 136 <213> ORGANISM: Artificial Sequence
 137 <220> FEATURE:
 138 <223> OTHER INFORMATION: Description of Artificial Sequence:linker
 139 <400> SEQUENCE: 7
 140 Gly Gly Arg Arg
 141 1
 143 <210> SEQ ID NO: 8
 144 <211> LENGTH: 5
 145 <212> TYPE: PRT
 146 <213> ORGANISM: Artificial Sequence
 147 <220> FEATURE:
 148 <223> OTHER INFORMATION: Description of Artificial Sequence:linker
 149 <400> SEQUENCE: 8
 150 Gly Gly Gly Gly Ser
 151 1 5
 153 <210> SEQ ID NO: 9
 154 <211> LENGTH: 8

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/897,844A

DATE: 11/23/2001
TIME: 10:57:17

Input Set : N:\Crf3\RULE60\09897844A.RAW
Output Set: N:\CRF3\11232001\I897844A.raw

```

155 <212> TYPE: PRT
156 <213> ORGANISM: Artificial Sequence
157 <220> FEATURE:
158 <223> OTHER INFORMATION: Description of Artificial Sequence:linker
159 <400> SEQUENCE: 9
160     Gly Gly Arg Arg Gly Gly Gly Ser
161         1             5
163 <210> SEQ ID NO: 10
164 <211> LENGTH: 9
165 <212> TYPE: PRT
166 <213> ORGANISM: Artificial Sequence
167 <220> FEATURE:
168 <223> OTHER INFORMATION: Description of Artificial Sequence:linker
169 <400> SEQUENCE: 10
170     Leu Arg Gln Arg Asp Gly Glu Arg Pro
171         1             5
173 <210> SEQ ID NO: 11
174 <211> LENGTH: 12
175 <212> TYPE: PRT
176 <213> ORGANISM: Artificial Sequence
177 <220> FEATURE:
178 <223> OTHER INFORMATION: Description of Artificial Sequence:linker
179 <400> SEQUENCE: 11
180     Leu Arg Gln Lys Asp Gly Gly Gly Ser Glu Arg Pro
181         1             5             10
183 <210> SEQ ID NO: 12
184 <211> LENGTH: 16
185 <212> TYPE: PRT
186 <213> ORGANISM: Artificial Sequence
187 <220> FEATURE:
188 <223> OTHER INFORMATION: Description of Artificial Sequence:linker
189 <400> SEQUENCE: 12
190     Leu Arg Gln Lys Asp Gly Gly Gly Ser Gly Gly Gly Ser Glu Arg Pro
191         1             5             10             15
193 <210> SEQ ID NO: 13
194 <211> LENGTH: 25
195 <212> TYPE: DNA
196 <213> ORGANISM: Artificial Sequence
197 <220> FEATURE:
198 <223> OTHER INFORMATION: Description of Artificial Sequence:ZFP target site
199     region surrounding initiation site of vascular
200     endothelial growth factor (VEGF) gene containing
201     two 9-base pair target sites
202 <220> FEATURE:
203 <221> NAME/KEY: protein_bind
204 <222> LOCATION: (4)..(12)
205 <223> OTHER INFORMATION: upstream 9-base pair ZFP VEGF1 target site
206 <220> FEATURE:
207 <221> NAME/KEY: protein_bind

```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/897,844A

DATE: 11/23/2001
TIME: 10:57:17

Input Set : N:\Crif3\RULE60\09897844A.RAW
Output Set: N:\CRF3\11232001\I897844A.raw

```

208 <222> LOCATION: (14)..(22)
209 <223> OTHER INFORMATION: downstream 9-base pair ZFP VEGF3a target site
210 <400> SEQUENCE: 13
211      agcggggagg atcgcgagg cttgg
213 <210> SEQ ID NO: 14
214 <211> LENGTH: 298
215 <212> TYPE: DNA
216 <213> ORGANISM: Artificial Sequence
217 <220> FEATURE:
218 <223> OTHER INFORMATION: Description of Artificial Sequence:VEGF1 ZFP
219      construct targeting upstream 9-base pair target
220      site in VEGF promoter
221 <220> FEATURE:
222 <221> NAME/KEY: CDS
223 <222> LOCATION: (2)..(298)
224 <223> OTHER INFORMATION: VEGF1
225 <400> SEQUENCE: 14
226      g gta ccc ata cct ggc aag aag aag cag cac atc tgc cac atc cag ggc 49
227      Val Pro Ile Pro Gly Lys Lys Lys Gln His Ile Cys His Ile Gln Gly
228      1          5          10          15
229      tgt ggt aaa gtt tac ggc aca acc tca aat ctg cgt cgt cac ctg cgc 97
230      Cys Gly Lys Val Tyr Gly Thr Thr Ser Asn Leu Arg Arg His Leu Arg
231      20          25          30
232      tgg cac acc ggc gag agg cct ttc atg tgt acc tgg tcc tac tgt ggt 145
233      Trp His Thr Gly Glu Arg Pro Phe Met Cys Thr Trp Ser Tyr Cys Gly
234      35          40          45
235      aaa cgc ttc acc cgt tcg tca aac ctg cag cgt cac aag cgt acc cac 193
236      Lys Arg Phe Thr Arg Ser Ser Asn Leu Gln Arg His Lys Arg Thr His
237      50          55          60
238      acc ggt gag aag aaa ttt gct tgc ccg gag tgt ccg aag cgc ttc atg 241
239      Thr Gly Glu Lys Lys Phe Ala Cys Pro Glu Cys Pro Lys Arg Phe Met
240      65          70          75          80
241      cgt agt gac cac ctg tcc cgt cac atc aag acc cac cag aat aag aag 289
242      Arg Ser Asp His Leu Ser Arg His Ile Lys Thr His Gln Asn Lys Lys
243      85          90          95
244      ggt gga tcc
245      Gly Gly Ser
247 <210> SEQ ID NO: 15
248 <211> LENGTH: 99
249 <212> TYPE: PRT
250 <213> ORGANISM: Artificial Sequence
251 <220> FEATURE:
252 <223> OTHER INFORMATION: Description of Artificial Sequence:VEGF1 ZFP
253      construct targeting upstream 9-base pair target
254      site in VEGF promoter
255 <400> SEQUENCE: 15
256      Val Pro Ile Pro Gly Lys Lys Lys Gln His Ile Cys His Ile Gln Gly
257      1          5          10          15
258      Cys Gly Lys Val Tyr Gly Thr Thr Ser Asn Leu Arg Arg His Leu Arg

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/897,844A

DATE: 11/23/2001

TIME: 10:57:18

Input Set : N:\Crf3\RULE60\09897844A.RAW

Output Set: N:\CRF3\11232001\I897844A.raw

L:45 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1
L:47 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1
L:79 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2
L:101 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3